

VU Research Portal

Effects of quality of implementation on the learning outcomes of secondary education mathematics curriculum

Hoek, D.J.; Terwel, J.; van Hout-Wolters, B.; van den Eeden, P.

published in

Curriculum and Teaching
2003

DOI (link to publisher)

[10.7459/ct/18.1.02](https://doi.org/10.7459/ct/18.1.02)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Hoek, D. J., Terwel, J., van Hout-Wolters, B., & van den Eeden, P. (2003). Effects of quality of implementation on the learning outcomes of secondary education mathematics curriculum. *Curriculum and Teaching*, 18(1), 5-25. <https://doi.org/10.7459/ct/18.1.02>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Effects of Quality of Implementation on the Learning Outcomes of Secondary Education Mathematics Curriculum

Dirk J. Hoek

Leiden University

Jan Terwel

University of Amsterdam and Vrije Universiteit Amsterdam

Bernadette H.A.M. van Hout-Wolters

University of Amsterdam

Pieter J.W.M. Van den Eeden

Vrij Universiteit Amsterdam

Abstract

This article reports on the effects of the quality of implementation of an experimental program on the learning outcomes of individual students. The experimental program involved working in cooperative groups and an instruction to use social and cognitive strategies. The central question was: can between-class differences in the quality of implementation explain effects on the learning outcomes of individual students? The results showed the more the curriculum was implemented according to the intentions, the larger the differences were in learning outcomes between low- and high-achieving students. The high-achieving students profited more from the high quality implementation than did the low-achieving students. Implications for curriculum theory and practice are discussed in the light of various perspectives on curriculum implementation.

Keywords: cognitive strategies, learning outcomes, cooperative groups, mathematics curriculum, curriculum implementation

Introduction

At the start of the school year 1993-94 a major innovation was introduced for the first stage of Dutch secondary education. All students were taught according to a national curriculum (a common curriculum called 'basic education'). This new curriculum contains com-

mon objectives for 15 subjects to be covered in two years or more. In the next stage students were allocated to different educational levels.

In the common curriculum new subjects, aims and classroom processes were formulated. For mathematics education these processes consist of learning in real life contexts and problem solving in co-operative groups. Students need to acquire social and cognitive competencies. In line with a constructivist approach, students are expected to play a more active part in the learning process (Terwel, 1999).

Recently Roelofs and Terwel (1999) concluded that the development and the implementation of this innovative curriculum are still far behind the expectations of educators and policy makers. The present study of the implementation of a program in secondary mathematics education can be seen in this context.

As the starting point for the present implementation study it is necessary to refer to a recent study by Hoek, Terwel and Van den Eeden (1999). They reported on a study involving an experimental program for small group learning. In this experimental program students were instructed to use social and cognitive strategies. It was found that high-achieving students profited more from working in small groups than low-achieving students did. This was not expected because this experimental program was developed to give low-achieving students the opportunity to benefit more from group work. The quality of implementation of the experimental program might have caused this effect. It is not our aim to recapitulate and prove again that high-achieving students profit more from group work than low-achieving students do. However, our interest will be directed towards the effect of the implementation of the experimental program as experienced by the students.

Theoretical Background

Research on implementation shows that even when the curriculum as a document is the same, teachers implement curricula differently (Fullan, 1991; Fullan & Pomfret, 1977; Snyder, Bolin & Zumwalt, 1992). It is to be expected that these differences in implementation will lead to different learning outcomes. Low and high levels of implementation will effect the social and instructional processes in the class.

To evaluate curricula it is important to know how the teacher deals with the formal curriculum within the educational setting (Goodlad, Klein & Tye, 1979; Snyder, Bolin & Zumwalt, 1992). There are at least three possible perspectives in the study of curriculum implementa-

tion: (i) the fidelity perspective, (ii) the mutual adaptation perspective and (iii) the enactment perspective (Snyder, Bolin & Zumwalt, 1992). It might seem attractive to regard these three perspectives as a continuum. In such a continuum mutual adaptation could be seen as a midpoint in a bipolar situation, in which complete implementation (fidelity) and enactment are the opposite poles. Although there is in fact some truth in this kind of representation, strictly speaking there is little logic in it. The opposite of fidelity is non-fidelity; the latter term indicating that true implementation is or has been non-existent. In the light of this, we prefer to discuss the three perspectives in their own right, but from a multidimensional point of view.

In all perspectives a curriculum is used to create *learning experiences* for students, and teachers are viewed primarily as the agents for the construction of learning experiences. The perspectives differ in their view on the role of teachers and students in constructing learning experiences. The perspectives also differ in their definition and their focus on particular representations of the curriculum.

Central to the fidelity perspective is the question of the degree to which a particular curriculum-on-paper is implemented as planned. Fidelity can be expressed as a percentage, ranging from 100 percent, indicating complete implementation, to 0 percent, indicating no implementation at all. The mutual adaptation perspective recognises that complete implementation is impossible in practice; a curriculum will always be adapted to a specific situation. Some will see adaptation as a kind of pragmatic compromise; others will regard it as essential or desirable. In both the fidelity and mutual adaptation perspectives the written curriculum materials and the corresponding plans provide the main focus of attention and they are compared to the operational and experiential curriculum.

More specifically, in the fidelity perspective the teacher's task is to do what is prescribed in the curriculum and implement the curriculum as intended regardless of situational differences. In the fidelity perspective the curriculum tends to be defined in terms of planned and consciously directed learning experiences. Without representations of the ideal and formal curriculum the fidelity perspective is not possible; operational and experiential representations of the curriculum are only relevant in so far as they are compared with the ideal and formal representations.

The mutual adaptation perspective has more regard for the responsibilities of the teachers and for situational differences. In the mutual adaptation perspective the curriculum also tends to be defined in

terms of consciously directed experiences. Simultaneously however, the impossibility of planning all relevant learning experiences for students in a formal curriculum is recognised on the one hand, while on the other hand the teachers are stimulated to adapt the conscious directions in the curriculum to real and concrete situations. In so doing teachers are stimulated to construct operational and experiential representations of the curriculum that differ from the ideal and formal ones. Real-life situations in the classrooms are always more complicated than can be anticipated in a formal representation. Mutual adaptation is, however, still a perspective in which educational administrators and implementation researchers are especially interested in the directed experiences of the students, and therefore compare the operational with the formal representations of the curriculum.

In the enactment perspective on the other hand, the curriculum materials are seen as tools for students and teachers, to be used in the construction of enacted classroom experiences. A different view of the function of curriculum is involved here. From the enactment perspective a curriculum is viewed as a set of educational experiences jointly created by students and teachers (Snyder, Bolin & Zumwalt, 1992). The enactment perspective stresses the creative role and full responsibilities of the teachers and the uniqueness of each situation. In this perspective the curriculum tends to be defined in terms of the whole range of students' experiences, both the directed and the undirected. The enactment perspective adheres to a constructivist approach while focusing on the operational and experiential representations of the curriculum. An enacted curriculum is exclusively an operational and experiential curriculum; it is not a formally documented one. The differences between the enactment perspective on the one hand, and the fidelity and adaptation perspectives on the other hand, concern primarily the function of the curriculum as a document. The differences have little to do with the perceived, the operational and experiential curriculum.

Studies on curriculum implementation have focused almost exclusively on the fidelity and mutual adaptation perspectives and consequently on the behaviour of the participating teachers.

External observers check and sometimes discuss the instruction given by the teachers. The observations are aimed at observing the teacher's behaviour and the way the teacher treats the participating students (Van den Akker, 1988; Van den Berg, 1996, 1997; Roes, 1997).

As the observation is aimed at teachers, this means that the students are treated as passive recipients of curricular innovations. Cheung, Hattie, Bucat and Douglas (1996) have discussed the possibility of using students as observers of their own learning environment, since students are the consumers of the curriculum and they experience the benefits and the problems of the curriculum implemented by their teacher. Students continuously observe and experience their teacher's behaviour as well. They are in touch with the teaching materials and the instruction given by their teacher. As a result, they can provide relevant information on curriculum implementation, because they have experienced the learning environment during the whole school year.

Since we have argued that in all perspectives on implementation, a curriculum is basically intended to provide *learning experiences*, it is important to know how these experiences can be investigated. There are methods such as direct observations and interviews. These two methods are expensive and consequently not often used in large-scale implementation research. However there is an alternative for obtaining an insight into the learning experiences of students. It is possible to use students as observers of the given curriculum by administering a questionnaire to them. This is an effective and reliable way to collect data from a large group of students (Taylor, Fraser & Fisher, 1997; Terwel, 1990).

Besides the general effect of implementation on learning outcomes, it is known from several studies that students with different abilities benefit differently from cooperative learning environments: high-achieving students profit more than low-achieving students do. This effect is known as the differential effect (see Hoek, Terwel & Van den Eeden, 1997; Leechor, 1988; Van den Eeden & Terwel, 1994; Webb, 1989, 1991). This article deals with the effect of the quality of implementation on learning outcomes in cooperative learning. In addition to the general effects of implementation there is an important question concerning differential effects. The question is who benefits more from a successful implementation: the high achievers or the low-achieving students? In a study by Van den Eeden and Terwel (1994) the following hypothesis was put to the test: the better the implementation of the curriculum, the higher the scores on the post-test. From this study it was concluded that the majority of students profited from the quality of implementation. In addition it was found that low-and medium-achieving students benefited most from the quality of implementation while the high-achieving students were less sensitive to differences in implementation (Van den Eeden & Terwel (1994: 462, 472).

Research Question and Hypothesis

The main research question addressed in this article is: what are the effects of between-class differences in the implementation of an experimental program for learning in cooperative groups with instructions to use social and cognitive strategies, on the learning outcomes of low- and high-achieving students in secondary mathematics?

The quality implementation hypothesis

The quality implementation hypothesis states that the more a curriculum is delivered in accordance with its implementation protocol, the better the learning outcomes will be. It was expected that the quality of implementation would have a positive effect on the learning outcomes of the individual students.

Method

Design

The research made use of a pre-test post-test design. In the present study the focus is on the quality of implementation of the experimental program. This program covered two series of 14 lessons in mathematics in the first year of secondary education.

Participants

The participants in the original study were 222 students (aged 12 to 13 years) comprising 9 classes at two Dutch secondary schools, with heterogeneous classes.

Assessments

The learning outcomes of the individual students were measured with a Mathematical Reasoning Ability Test. This test was administered to all participating students prior to and at the end of the experiment (Horn, 1969). According to Horn (1969) this test leans strongly on the factor 'reasoning'. In earlier research the correlation between Mathematical Reasoning Ability Test and mathematical achievement varied between .50 and .80 (Aurin, 1966; Terwel, Herfs, Mertens & Perrenet, 1994). Scores on the pre-test were used as a control variable, with the post-test scores as the dependent variable. The total test contains 80 items, with possible scores ranging from 0 to 80 correct items.

To measure the quality of the implementation of the experimental program a questionnaire was administered immediately after the experimental period. The questionnaire was a curriculum-specific version of the learning environment scale (Fraser & Tobin, 1989; Taylor, Fraser & Fisher, 1997). This questionnaire contains items about the teacher's instruction and guidance, cooperation between

students, and instruction strategy. It contains statements on which the student could mark a digit from 1 ("hardly ever") to 5 ("nearly always").

Procedure and Materials

The experiment started with a teacher-training program. In this program the basic instruction (AGO) model used in the experiment was explained to all the participating teachers. AGO is a Dutch acronym for 'adaptive instruction and cooperative learning'. The AGO model was used as the point of departure in the design of the experimental and control program. This instructional model combines aspects of cooperative learning and adaptive instruction. The model is based on assumptions about cooperative learning and cognitive learning theory. The AGO model is designed for the middle grades, 12- to 16-year-old students (Freudenthal, 1973; Terwel, 1990; Terwel, Herfs, Mertens & Perrenet, 1994). In terms of Mason and Good's (1993) classification, the model can be described as a whole-class model that allows for student diversity through small group ad hoc remediation and enrichment. The AGO-model consists of the following stages: 1. whole-class introduction of a mathematics topic in real-life contexts; 2. small group co-operation in heterogeneous groups of four students; 3. teacher assessments: diagnostic test and observations; 4. alternative learning tracks depending on assessments.

These tracks consist of two different modes of activity: 4a. individual work at student's own pace and level (enrichment) in heterogeneous groups with the possibility of consulting other students; 4b. opportunity to work in a remedial group (scaffolding) under direct guidance and supervision of the teacher; 5. individual work at student's own level in heterogeneous groups with opportunities for students to help each other; 6. whole-class reflection and evaluation of the topic; 7. a final test. The AGO model was the basic instructional design for the experimental program. However, some modifications were made, especially in stages 4 and 5.

In the experimental program, heterogeneity, as in the original AGO model, was abandoned during stages 4 and 5, and replaced by homogeneous small groups. The teachers who, for practical reasons, preferred a kind of within-class ability grouping (within-class setting with low-, medium-, and high-achieving student groups) in stages 4 and 5 advocated this modification. Rearrangement of students was thus no longer required during these stages. In addition, the teacher could locate low-achieving students more easily, because they were seated in the same ability group.

After the first training session for teachers there was another session during which the teachers practised the instruction of social and cognitive strategies they were to give to their students. For the teacher-training program a manual was developed, containing the following elements: (1) the design of the experimental program, with the student assessments, (2) the content of the domain-specific topics, (3) the instruction model for the use of social and cognitive strategies; and (4) exercises for the student training program.

Before the students started the program their teachers trained them. This training consisted a whole-class instruction to demonstrate the problem-solving strategy to be used during the experimental period (van Streun, 1989, 1994). During this training the teacher served as a leader or expert for his students in a cognitive apprenticeship model. The methods used involved exploration, coaching, articulation, and reflection (Collins, Brown & Newman, 1989; Schoenfeld, 1985; 1992). In the second part of the students' training the teacher trained them in cooperative groups using the 'Broken Circles Problem' and the 'Master Designer Problem' (Cohen, 1986: 159-164). These problems were used to discuss some rules for small group work. After working in small groups on each 'problem' the teacher organised a class discussion. During this discussion the students considered the significance of the problem and what they had learned. At the end of the student training students were given rules for communication, practice, and feedback concerning small group cooperation. These rules were put on a big poster, which was fixed on the classroom wall.

Two series of lessons were developed for the experiment. The first series of lessons ('Measures and Measurement') consisted of problems about measurement units, circumference and area. The second series of lessons ('Information Gathering') consisted of problems about tables and figures with different kinds of information, maps and graph representations. Before and after the experimental period assessments were carried out. The experimental booklets contained mathematical problems with extra assignments for problem solving and extra assignments to let students work effectively in cooperative groups.

Results

Data and analysis at the individual level

The reliability coefficient (r) of the Mathematical Reasoning Ability Test for the pre-test and post-test was .80. A paired t-test showed a significant difference between the pre-test ($M=51.65$, $SD= 9.24$) and post-test ($M=53.5$, $SD= 8.14$) [$t(221) = 3.5$, $p < .05$]. Now we turn to

the analysis of the implementation data. The Implementation Questionnaire (IML) had a reliability coefficient α of .92. Table 1 (for all tables and figures, see appendix) shows the mean scores of the classes in the experimental program on the quality of implementation as perceived by the classes.

From an analysis of variance of the data in Table 1 it can be concluded that there were significant differences between the classes [$F(8, 213) = 21.38, p < .05$]. To answer the main research question a multilevel analysis was used. The analysis focuses on the effects of the implementation at class level, using the Implementation questionnaire on the learning outcomes of the student.

The multilevel analysis

To assess the effects of the program with regard to the quality of implementation, the appropriate tool to use was a multilevel model of analysis (Van den Eeden & Terwel, 1994). This gives the possibility of analysing for differential effects. The M3-E statistical program was used in the multilevel analysis (Posser, Rossbach & Goldstein, 1993). The mathematical reasoning ability of the student and the quality of implementation as perceived by the class formed continuous variables. The analysis was directed at the relation between the quality of implementation, the mathematical reasoning ability of the student before the experimental period (pre-test), and the mathematical reasoning ability after the experimental period (post-test). In the model used in the analysis these relations are described in terms of regressions of a dependent variable on the independent variables at the student level and at the class level.

Figure 1 represents the multilevel model tested in this study. Path 1 represents the effect of the pre-test (pre-knowledge) on the post-test. Path 2 represents the differential effect: a positive sign for the coefficient means the high-achieving students benefit more from the program than low-achieving students. Path 3 represents the quality implementation hypothesis, a positive sign for the coefficient means that the better the teacher implemented the program, the more the high-achieving students benefited from the experimental program.

The results of the multilevel analysis

Table 2 shows the results of the multilevel analysis for the implementation of the experimental program. Table 2 represents the first model (model 1) and the last model (model 2) of the multilevel analyses; the intervening models are included in this table. In Table 2 Model 1, the variance at the individual level (within classes) was 60.3 or 97% of the total variance, while the variance at the class level

(between classes) is 2.0, or 3% of the total variance. In Model 2 the differential effect and the implementation effect on the differential effect were introduced. The differential effect is significant and positive (0.006), meaning that the difference in learning outcomes between low- and high-achieving students increased with the quality of implementation, with high-achieving students profiting more than the low-achieving students did. In addition, the quality of implementation of the experimental had a significant and positive effect on the differential effect (0.0002); implementation intensified the difference between the low- and high-achieving students.

In Figure 2, Model 2 is represented. Figure 3 displays the results of two regressions for classes with the lowest and highest score on quality of implementation according to Model 2 in Table 2. Figure 3 plots the graphs of two classes: the class with the highest and the lowest score on the quality of implementation. To plot these graphs we used the model represented in Figure 2. It should be borne in mind that the scores of the students of the other classes fall between these maximum and minimum lines.

Figure 3 shows that there is a learning gain for all students. The learning gain of the low- and high-achieving students differs. To interpret Figure 3 we have to look at the relative learning gain of the low- and high-achieving students. This relative learning gain is operationalised by the gradient of the implementation graphs. The gradient of the both graphs becomes larger per unit when the score on the pre-test is higher. So, the higher the student scores on the pre-test, the bigger the relative learning gain is. This means that the high-achieving students profited more from the experimental program than the low-achieving students did. Figure 3 shows that the better the quality of implementation, the stronger the effect is. The regression graph of high implementation is steeper than the regression graph of the low implementation. High quality of implementation of the experimental program leads to a higher relative learning gain of the high-achieving students, while the relative learning gain of the low-achieving students of the two programs does not differ much.

Conclusions and Discussion

In this study we were looking for the effects of differences in implementation of an experimental program. From the literature it is known that implementation is always a matter of degree. In the present study substantial and significant differences in implementation between classes were found. What were, in general, the effects of these differences on the learning outcomes of the participating students? Did low- and high-achieving students differentially benefit

from the implementation? It turned out that the better the quality of the implementation of the experimental program the bigger the learning gain of the students was. In addition it appeared that the high-achieving students were the ones who profited most from the implementation. These outcomes confirm partly the study of Van den Eeden and Terwel (1994) as far as the majority of the students are concerned. As expected, most students profited from a higher quality of implementation. However the outcomes concerning the differential effects were in contradiction to the earlier study. In the study of Van den Eeden and Terwel (1994) the low- and medium-achieving students were most stimulated by the quality of implementation while in the present study the high achievers profited most. How can these differences be explained? Between the two studies there was an important difference in the grouping practice. In the study of Van den Eeden and Terwel (1994) the cooperative groups were heterogeneous according to the 'Freudenthal model', while in the present study teachers explicitly asked for homogeneous 'within class grouping' for practical reasons. Therefore we will discuss in some detail the possible factors and implications of the outcomes of the present study, also in the light of what is known from other studies.

Even when the teachers give instruction in which they take into account the different levels of the students and pay extra attention to the low-achieving students, the instruction is more beneficial for the high-achieving students: they profit the most from the instruction. The extra attention given to the low-achieving students did not counterbalance or mitigate the differences in learning outcomes between the low and high-achieving students.

A possible explanation can be found in the structural complexity of the intervention for both the teachers and the students. The teachers were used to give whole-class instruction according to the 'information transmission model' (Greeno, 1991). Within this model a teacher explains how you have to solve a problem, sometimes explaining why they solved it the way they did (information transmission model). However, within this experiment they had to ask and answer questions, which taught students to reflect on the problem-solving process and procedures. On the other hand, students are not used to giving an answer and providing an explanation on "why" or "how" questions. In particular, the low-achieving students have difficulties in doing this. The low-achieving students have problems understanding that it is important to ask these questions. The low-achieving students want their teacher to explain how to solve a certain type of problems. In addition to the use of problem-solving strategies, the teachers instructed and guided the students to use strategies for effective coop-

erative groups. This makes the instructions even more complex, maybe too complex for both the teachers and the low-achieving students. When the teacher simply provides a model of how to solve a problem it is questionable whether students will obtain an understanding of mathematics and be able to solve new (different kinds of) problems. Implementing a rather complex instructional program might not be a wise idea if time constraints are rather tight. Given the complexity the teachers think this is too difficult to implement. This leaves implicitly room to adapt the instructional program to their 'normal' way of handling. When the implementation of a program is too difficult a teacher will be less eager to implement the program, in the way the researchers want them to do.

In order to understand the outcomes of this study we have to look at both the structural complexities of the intervention and the outcome assessment. In terms of the classification of Hattie, Biggs and Purdie (1996) the present intervention can be described as a complex intervention (multi-structural/relational) with an assessment type less closely related to the intervention (far transfer). From the meta-analysis of Hattie, Biggs and Purdie (1996) it is to be concluded that especially medium- and high-achieving students benefit from this kind of interventions. In addition Hattie, Biggs and Purdie conclude that low-achieving students seem to be unable to benefit from interventions of most kinds with a few exceptions.

Another explanation might be the homogeneous grouping after the diagnostic test. During this stage, students were put into homogeneous ability groups. Students in homogeneous low ability groups are working in a poorer learning environment than high-achieving students. According to Dar and Resh (1986a, 1986b) low-achieving students are more sensitive to the quality of their learning environment. High-achieving students have more resources of their own. As a consequence high achievers are less vulnerable or dependent on resources in their learning environment. Ability grouping is not a zero sum game in the sense that the loss of low-achieving students is compensated by the gain of high-achieving students (if 'compensation' could apply here). Anyway, the 'differential sensitivity hypothesis' explains why low-achieving students suffer from homogeneous grouping practices whether these are streaming, setting or within class ability grouping (see also Hallinan, 1987).

During this stage the teacher coached the low-achieving students. However, he also had to coach other students for a great part of the lessons. During this coaching the teacher was the only rich resource in the group of low-ability students. Therefore resources were restrict-

ed and might have resulted in the learning gain of the low-ability students being smaller than the learning gain of the high-ability students.

Finally we turn to the implications for theory and practice of curriculum implementation. In curriculum literature it is implicitly or explicitly assumed that the quality of implementation positively affects the learning outcomes of students. This assumption was confirmed in the present study. However, to the best of our knowledge, the question of how the quality of implementation differentially affects the learning gains of high and low-achieving students remains practically unanswered. In the present study we found a phenomenon which is often found in educational innovations: the rich are getting richer. For example the results from implementation of a peer-tutoring program go in the same direction. All students profited but the chief beneficiaries appear to be the most advanced students. 'In the past, advanced students progressed exponentially by volunteering as tutors for underachievers. The rich got richer as they learned and mastered material by taking the responsibility of teaching it to others (Gartner & Riesman, 1998). There is nothing wrong with exponential growth for the brightest students as long as there are no detrimental effects for the poorer-achieving students. The key would appear to lie in the learning experiences in the cooperative groups. In our study the more the curriculum was implemented according to the (cooperative) intentions the more the high-achieving students got the opportunity to articulate their ideas and to confront and share their views with others in a rich learning environment. Although in principle the same curricular opportunities were available for the low-achieving students, in practice their learning environment was impoverished as a consequence of the homogeneous grouping practice. This grouping practice seems especially detrimental for low-achieving students in situations where higher order thinking and strategic learning is involved. 'Matthew effects' in curriculum innovations are difficult to avoid (Hallinan, 1987; Dar & Resh, 1994; Gartner & Riesman, 1998).

In the theoretical background section we described three perspectives on curriculum implementation. In the fidelity perspective the formal curriculum is used to determine the discrepancy between the intentions and practices. The 'adaptation' and 'enactment perspectives' stress the creative role and full responsibilities of the teachers. The results of this study show that teachers indeed can propose and even demand curricular adaptations e.g. concerning the grouping practice. These adaptations may be in contradiction with with basic characteristics of the innovation, and may seem not always to be effective from a researcher's point of view. 'Enactment' and 'Constructivist'

perspectives on implementation offer a significant challenge for curriculum theory and practice, in that it can break down the harness of the 'fidelity' perspective and permit teachers and students to construct their own learning environment. However at the same time the empowerment of teachers and students creates room for the delusions of the day. The danger of fashionable practices can be avoided if the relevant research literature and the perceptions of the students are taken in consideration. Therefore the solution lies in the consultation processes between teachers, curriculum designers and researchers. Student perceptions of the curriculum as described in this article can be used as an input to this consultation process.

References

- Aurin, K. (1966) *Bildung in neuer sicht, ermittlung und erschliessung von begabungen im ländlichen raum* [Education in a new perspective, coaching and opening up the giftedness in the different federal states]. Villingen, Germany: Neckar-Verlag.
- Cheung, D., Hattie, J., Bucat, R., & Douglas G. (1996) Students' perception of curriculum implementation, *Curriculum and Teaching*, 11, 49-60.
- Cohen, E. G. (1986) Designing groupwork: strategies for the heterogeneous classroom. Stanford, CA: Stanford University.
- Collins, A., Brown, B. S., & Newman, S. E. (1989) Cognitive apprenticeship: Teaching the crafts of reading, writing and mathematics. In L. B. Resnick (ed.), *Knowledge, learning and instruction. Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Erlbaum.
- Dar, Y., & Resh, N. (1986a) Classroom intellectual composition and academic achievement, *American Educational Research Journal*, 23, 357-374.
- Dar, Y., & Resh, N. (1986b) *Classroom composition and pupil achievement: Study of the effects of ability based classes*. London, England: Gordon and Breach.
- Dar, Y., & Resh, N. (1994) Separating and mixing students for learning: concepts and research, *Pedagogisch Tijdschrift*, 19, 109-126.
- Fraser, B. J., & Tobin, K. (1989) Combining qualitative and quantitative methods of the study of classroom learning environment. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Freudenthal, H. (1973) *Mathematics as an educational task*. Dordrecht, The Netherlands: Reidel.
- Fullan, M. (1991) Curriculum implementation. In A. Lewy (ed.), *The international encyclopedia of curriculum* (pp. 378-384). Oxford, England: Pergamon Press.
- Fullan, M., & Pomfret, A. (1977) Review of research on curriculum and instruction implementation, *Review of Educational Research*, 47, 335-397.
- Gartner, A. & Riesman, F. (1998) Turning Peer Pressure Inside Out, *Education Week*, 36, May 20, 1998.
- Goodlad, J. J., Klein, M. F., & Tye, K. A. (1979) The domains of curriculum and their study. In J. J. Goodlad et al. (Eds.), *Curriculum inquiry: The study of curriculum practice* (pp. 43-76). New York: McGraw-Hill.

- Greeno, J. G. (1991) Number sense as situated knowing in a conceptual domain, *Journal for Research in Mathematics Education*, 22, 170-218.
- Hallinan, M. T. (1987) Ability Grouping and Student Learning. In M. T. Hallinan (ed.), *The Social Organization of Schools: New Conceptualizations of the Learning Process* (pp. 41-67). New York/London: Plenum Press.
- Hattie, J., Biggs, J., & Purdie, N. (1996) Effects of learning skills interventions on student learning: a meta-analysis, *Review of Educational Research*, 36, 99-136.
- Hoek, D. J., Terwel, J., & Van den Eeden, P. (1997) The effects of social and cognitive strategies instruction on the mathematics achievement in secondary education, *Educational Research and Evaluation*, 3, 364-389.
- Hoek, D. J., Van den Eeden, P., & Terwel, J. (1999) The effects of integrated social and cognitive strategies instruction on the mathematics achievement in secondary education, *Learning and Instruction*, 9, 427-448.
- Horn, W. (1969). Prüfsystem für schul- und bildungsberatung (Handanweisung) [Testing system for school and developmental advice]. Göttingen, Germany: Verlag für Psychologie Hochrefe.
- Leechor, C. (1988) How high and low-achieving students differentially benefit from working together in cooperative small groups. Stanford, CA: Stanford University, School of Education. (unpublished dissertation).
- Mason, D. W., & Good, T. L. (1993) Effects of two-group and whole-class teaching on regrouped elementary students' mathematics achievement, *American Educational Research Journal*, 30, 328-360.
- Posser, F., Rossbach, J., & Goldstein, H. (1993) *ML-3E software for three level analysis: Users guide*. London: University of London, Institute of Education.
- Roelofs, E. & Terwel, J. (1999) Constructivism and authentic pedagogy: state of the art and recent developments in the Dutch national curriculum in secondary education, *Journal of Curriculum Studies*, 31, 201-227.
- Roes, M. (1997) *Nascholing op basis van lesvoorbeelden, in de context van curriculumvernieuwing* [In-service education based on lesson examples within the context of curriculum development]. Enschede, The Netherlands: PrintPartners Ipskamp. (dissertation).
- Schoenfeld, A. H. (1985) *Mathematical problem solving*. San Diego, CA: Academic Press.
- Schoenfeld, A. H. (1992) Learning to think mathematically: Problem solving, Metacognition, and sense making in mathematics. In G. A. Grouws (ed.), *Handbook of research on mathematics teaching and learning* (pp. 165-197). New York: MacMillan.
- Snyder, J., Bolin, F., & Zumwalt, K. (1992) Curriculum implementation. In Ph. W. Jackson (ed.), *Handbook of research on curriculum* (pp. 402-435). New York: Macmillan.
- Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997) Monitoring constructivist classroom learning environments, *International Journal of Education Research*, 27, 293-302.
- Terwel, J. (1990) Real mathematics in co-operative groups. In N. Davidson (ed.), *Co-operative learning in mathematics* (pp. 228-264). Menlo Park, CA: Addison-Wesley.
- Terwel, J. (1999) The challenge of constructivism for curriculum theory, practice and research, *Journal of Curriculum Studies*, 31, 195-200.
- Terwel, J., Herfs, P. G. P., Mertens, E. H. M., & Perrenet, J. Chr. (1994) Co-oper-

- ative learning and adaptive instruction in a mathematics curriculum, *Journal of Curriculum Studies* 26, 217-233.
- Van den Akker, J. J. H. (1988) *Ontwerp en implementatie van natuuronderwijs* [Design and implementation of science education]. Amsterdam, The Netherlands: Swets & Zeitlinger. (dissertation).
- Van den Berg, E. (1996) Effects of in-service education on implementation of elementary science. Enschede, The Netherlands: PrintPartners Ipskamp. (dissertation).
- Van den Berg, E. (1997) Nascholing natuuronderwijs: effecten op de attitude van leerkrachten [In service science education: Effects on the attitude of teachers]. *Tijdschrift voor de didactiek der β -wetenschappen*, 17, 30-55.
- Van den Eeden, P., & Terwel, J. (1994) Evaluation of a mathematics curriculum: differential effects, *Studies in Educational Evaluation*, 20, 457-475.
- Van Streun, A. (1989) *Heuristisch wiskunde-onderwijs: verslag van een onderwijs experiment* [Heuristic mathematics education: report of an educational experiment]. Groningen, The Netherlands: Rijksuniversiteit Groningen. (dissertation).
- Van Streun, A. (1994) Hoe onderwijs je probleemoplossen? [How do you teach problem solving?]. *Tijdschrift voor didactiek der β -wetenschappen*, 12, 210-225.
- Webb, N. M. (1989) Peer interaction and learning in small groups, *International Journal of Educational Research*, 1, 21-39.
- Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups, *Journal for Research in Mathematics Education*, 22, 366-389.

Appendix: The Jones family

After a whole-class introduction by the teacher, students work together in cooperative groups on mathematical problems in real-life contexts: Jones family. Below a sample of the curriculum material is given, followed by a protocol of the group problem solving process of the group. The group consists of two girls (Elly and Nicole) and two boys (Mark and Simon). The students do not differ much in their score on a mathematical reasoning ability. The group consists of the following four students (pre-test scores are also given):

Mathematical reasoning score (pre-test)

Simon	58
Mark	55
Elly	53
Nicole	56

Theo: the teacher

Samples of curriculum materials and observations in the cognitive condition

The Jones family car uses one litre of fuel for every 18 kilometres.

1. How many litres of fuel does the car use on a 180 kilometres journey?
2. The distance to the campsite in Hopetown is 900 kilometres. How many litres of fuel does the car use on that journey?

3. Fill in the blank in the following sentence: *The number of kilometres divided by equals the number of litres of fuel used by the car*
4. Fill the machine below.

Sample Protocols from Observations

Protocol problem 5a

- Elly Ten litres, because you *have to divide 18 kilometres by 180 kilometres, and that equals 10 kilometres* (the student makes the correct calculation but the result should be 10 litres rather than 10 kilometres).
- Simon The next problem is B (Simon wants to continue with the next problem).
Here one of the students is giving the wrong explanation. He probably means 180 divided by 10. He arrives at the correct result, but the unit used in the answer is incorrect.

Protocol problem 5b

- Mark *Yes, because five times 18 equals 90 and $50 \times 18 = 900$ (the student explains).*
- Simon That's right (Simon agrees).
Nicole immediately reads the following problem aloud.
One student explains a strategy to another student. He shows how easy it is for him to calculate in units of ten

Protocol Problem 5d

- Simon Does 1:18 equal 18 (Simon answers again).
Somebody else continues $18 \times 50 = 900$.
Mark (Elly asks Mark to continue). No, *wait. What do you have? (Elly wants to know what Mark's solution is). No, that IS right, because it is just an example (either Mark or Simon answers). Yes, because $18 \times 50 = 900$ (one of the girls reconfirms the solution). So it's 18. Mark has the correct solution in his exercise book but does not dare mention it. He accepts the wrong solution given by the others as correct.*
- Simon Is that all right?
Theo come and have a look. Simon explains to Theo what's going on.
- Simon Yes, Theo look! She has a completely different solution to what she has. The solution to Problem C was 18, you see.
- Theo The number of kilometres divided by 18 equals the number of litres of fuel that are used (Theo explains). Fill the machine.
- Theo Do you put 1 in, or do you put something else in (Theo asks Simon).
- Simon No, something else (Simon correctly answers Theo's question whether he should fill in 1, or something else).
- Simon You put it in here, the number of kilometres. You don't only have to put in the number 1. You can put in any number. Yes, you can agree what to put in here; but what do the rest have?
- Simon Uh
- Elly $18 \times 50 = 900$ (Elly answers)
- Simon Yes, that's what I had as well (Simon confirms)

- Theo And why do you say 18×50 here? Why 18? Where do you get the 50 from?
(Theo asks the group).
The students return to problem B to explain where they get 50 from. Elly explains to Theo where she got the 50 from.
- Theo Does it show anywhere that you have to put 18 into the machine rather than anything else? (Theo asks the group).
- Elly No because it says 18 here, because you ... the number of kilometres divided by 18 is the number of litres that are needed (Elly answers).
- Theo So you need the number of litres of fuel that's needed? (Theo asks Elly).
Yes.
But it says, fill the machine below
Simon wants to press on and reads problem D aloud
- Theo OK (Theo reacts and continues his explanation). You start with the number of kilometres. What do I have to do with the number of kilometres to get the number of litres of fuel?
- Simon You divide it (Simon answers correctly).
- Theo That's right, you divide it by 18 (Theo confirms the correct answer). Is it a times machine or a dividing machine?
A dividing machine (one of the students answers).
- Theo It's a divide-by machine (Theo confirms).
- Mark You see! (Mark). *Mark had the correct solution but he didn't dare to say that he did not agree with the other students' solution. Divided by 18, right?*
- Mark (Mark asks Theo another time).
- Mark *So I was right after all (Mark). So you don't have to put 50 kilometres into it at all, because one family drive 50 kilometres, they cover 50 kilometres and the ... What's coming out now (Simon is getting impatient). What comes out of it is the number of litres, you see (Theo repeats the explanation). Oh, I see (Simon recognises Theo's explanation).*
- Theo What you put in this machine... (Theo wants to see if Simon has really understood, but Simon interrupts him).
- Simon Yes, I understand it (Simon). If I put 180 kilometres in the machine, and then the machine says you have to divide by 18 and the results is the number of litres. *But the number of kilometres that I put in the machine is not fixed.*
Theo once more repeats the explanation. There follows an inaudible discussion between Simon and Theo. Is it clear now that this is not a times machine but a divide-by machine? (Theo asks whether everybody has understood). And then you have to take the table on the next page, you have to change that to the result you got (Theo shows specifically how to change the next problem). So we divide by 18 (one of the students answers).
- Theo Yes, you divide by 18 (Theo confirms).

Reflection

In the italicised part of the protocol we find a continued mix of the 'sucker-effect' and status differential-effect (Salomon & Globerson, 1989). Mark, a rela-

tively high-achieving student adopts the wrong solution and thinks his own solution is wrong instead. Later on in the protocol Mark sees that he had given the correct solution after all. Another striking aspect of his behaviour is that he hardly takes part in the group interaction. After the lesson the observer asked Mark why he had not asked or said anything during the group-interaction. Mark answered: "Most people are better at maths than I am".

With regard to the structure of the problem, we conclude that this does allow for a concrete completion of the solution. In parts a. and b. questions are asked about how many litres of fuel a car uses to cover a certain distance. The students subsequently calculate the distance by dividing the number of kilometres by 18. In filling the machine they have to understand that, if they want to find out about fuel consumption by way of the number of kilometres, they have to divide the number of kilometres by 18. Filling in the verbal formula often presents no problems. However, converting the rule to a machine does. They stick to the concrete numbers. It is very difficult for the teacher to explain that the general rule for the problem in hand has to be stated in the machine. It remains to be seen whether the students have shown real understanding. In interaction processes there are social and cognitive factors involved.

Table 1: Scores at the class level on the post administration of the quality of implementation of the experimental program.

Class	M	SD
1	3.59	0.35
2	3.59	0.38
3	3.70	0.55
4	2.43	0.47
5	3.69	0.49
6	3.20	0.38
7	3.61	0.52
8	3.54	0.44
9	3.04	0.48

Table 2: The results of the multilevel analysis, with the score on the mathematical reasoning ability post-test as dependent variable (standard deviation between parentheses), N-students = 222, N-classes = 9.

	Model 1		Model 2	
	β	S.E. of β	β	S.E. of β
Fixed part				
Pre-test				
Pre-test-squared			0.006	0.003
Explaining between class slope differences pre-test-squared				
Implementation			0.0002	0.0001
Random Part				
σ^2 (student)	60.3	5.8	40.0	3.8
Class:				
σ_1^2	2.0	2.1	0.0	0.0
Model statistics				
Likelihood ratio	1545.3		1449.1	
Difference	-		96.2	
Difference degree of freedom	-		3	
Difference with model	-		1	

Figure 1: Model tested in the multilevel analysis

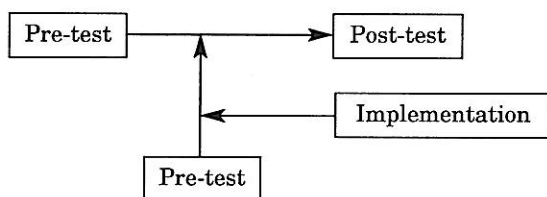


Figure 2: Results of the multilevel analyses (Table 2, model 2).

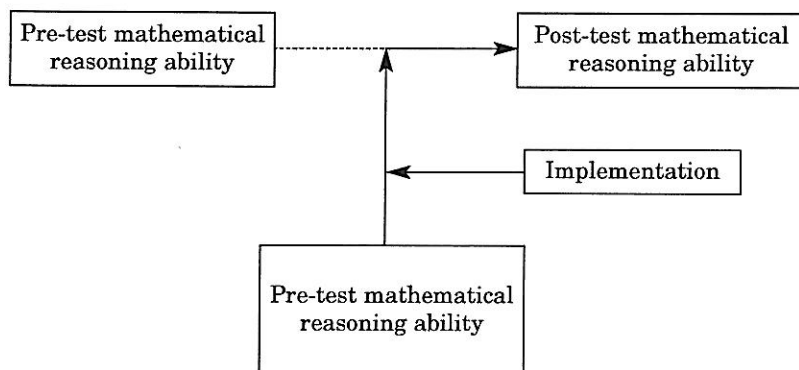


Figure 3: Graphs of the post-test on the pre-test, according to Model 2 for classes with the lowest and highest score on quality of implementation.

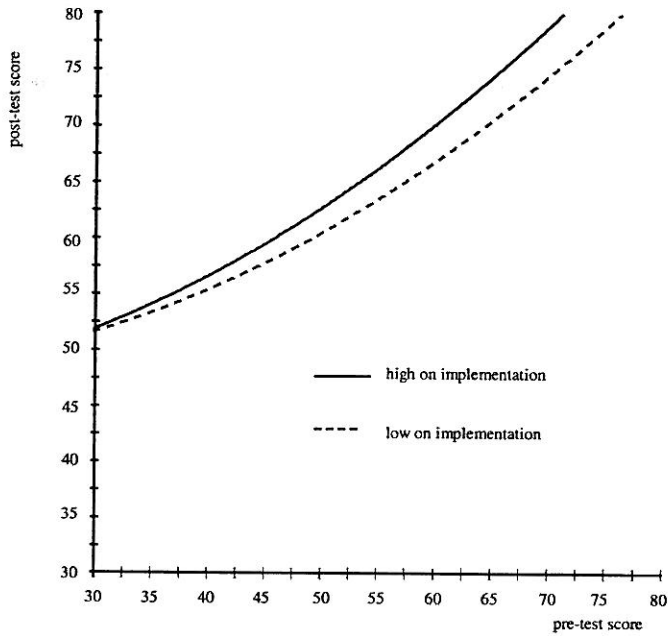


Figure 4: Machine of the problem

